DOCUMENT RESUME

ED 446 904 SE 063 736

AUTHOR Cartier, Jennifer

TITLE Assessment of Explanatory Models in Genetics: Insights into Students' Conceptions of Scientific Models. Research Report.

INSTITUTION Wisconsin Univ., Madison. National Center for Improving Student Learning and Achievement in Mathematics and Science.
SPONS AGENCY National Science Foundation, Arlington, VA.
REPORT NO NCISLA-98-1
PUB DATE 2000-01-00
NOTE 26p.
CONTRACT REC-9554193
AVAILABLE FROM National Center for Improving Student Learning and Achievement in Mathematics and Science, University of Wisconsin-Madison, 1025 W. Johnson Street, Madison, WI 53706; phone: 608-263-3605; e-mail: ncisla@education.wisc.edu. For full text: http://www.wcer.wisc.edu/ncisla
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Concept Formation; Evaluation; Genetics; High Schools; *Instructional Effectiveness; *Models; Problem Solving; Science Education; *Scientific Methodology; *Scientific Principles; *Teaching Methods

ABSTRACT

Past research in a high school genetics classroom has shown that groups of students solving computer-generated genetics problems differed in their application of assessment criteria to the tentative models they generated, and that consistent application of certain model-assessment criteria correlated with success in proposing adequate explanatory models. This paper describes a study in which classroom instruction in one high school class was focused on four specific model-assessment criteria: two involving fit between data and model, and two involving conceptual consistency, or fit within the model itself or between it and other models of scientific knowledge. Individual students' use of those assessment criteria during two round of model-revising problem solving was monitored and analyzed. All students were able to assess models based on empirical fit, but despite explicit instruction on their use, students did not consistently apply the conceptual assessment criteria to their models. This paper discusses the implications of these results on the use of modeling as an approach to teach students about the structure of scientific knowledge and the nature of science as a modeling activity. (Contains 19 references.) (Author)

Reproductions supplied by EDRS are the best that can be made from the original document.

ED 446 904

RESEARCH REPORT

NATIONAL CENTER FOR IMPROVING STUDENT LEARNING
AND ACHIEVEMENT IN MATHEMATICS AND SCIENCE

ASSESSMENT OF EXPLANATORY MODELS IN GENETICS:

Insights into Students' Conceptions of Scientific Models

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

S. Anderson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

JENNIFER CARTIER

University of Wisconsin-Madison

January 2000

Report No. 98-1

NCISLA · Wisconsin Center for Education Research · School of Education, University of Wisconsin-Madison · 1025 West Johnson Street · Madison, Wisconsin 53706
Phone: (608) 263-3605 · E-mail: ncisla@education.wisc.edu · Web site: http://www.wcer.wisc.edu/ncisla

SE 063 736

BEST COPY AVAILABLE

NATIONAL CENTER FOR IMPROVING STUDENT LEARNING
& ACHIEVEMENT IN MATHEMATICS & SCIENCE

NCISLA/Mathematics & Science
1025 W. Johnson Street
Phone: (608) 263-3605
ncisla@soemadison.wisc.edu

University of Wisconsin–Madison
Madison, WI 53706
Fax: (608) 263-3406
www.wcer.wisc.edu/ncisla

DIRECTOR: Thomas P. Carpenter

ASSOCIATE DIRECTOR: Richard Lehrer

COMMUNICATION DIRECTOR: Susan Smetzer Anderson

ABOUT THE CENTER

The National Center for Improving Student Learning & Achievement (NCISLA) in Mathematics & Science is a university-based research center focusing on K-12 mathematics and science education. Center researchers collaborate with schools and teachers to create and study instructional approaches that support and improve student understanding of mathematics and science. Through research and development, the Center seeks to identify new professional development models and ways that schools can support teacher professional development and student learning. The Center's work is funded in part by the U.S. Department of Education, Office of Educational Research and Improvement, the Wisconsin Center for Education Research at the University of Wisconsin-Madison, and other institutions.

SUPPORT

The preparation of this manuscript was supported, in part, by the National Science Foundation (Grant No. REC-9554193) and by the National Center for Improving Student Learning and Achievement in Mathematics and Science under the auspices of the Department of Education's Office of Educational Research and Improvement. Any opinions, findings, or conclusions are those of the authors and do not necessarily reflect the views of the supporting agencies.

ACKNOWLEDGEMENTS

I wish to thank the following for providing comments on the manuscript: Jim Stewart, Sam Donovan, John Rudolph, Cindy Passmore, and Sue Johnson. I also wish to thank John Rudolph for creating earlier versions of the Black Box description (Figure 1) and the representation of the Mendel simple dominance model (Table 1).

ABOUT THE AUTHOR

Dr. Jennifer Cartier is an assistant researcher at the NCISLA and an associate lecturer of Curriculum and Instruction at the University of Wisconsin–Madison. Her current research focuses on longitudinal studies of students' ability to create and justify arguments in science, and teacher professional development with respect to enhancing student argumentation.

Jennifer Cartier also works collaboratively with other Center researchers on The Modeling for Understanding in Science Education (MUSE) project. More information on this project will be posted at www.wcer.wisc.edu/ncisla in the Spring 2000.

BEST COPY AVAILABLE

ASSESSMENT OF EXPLANATORY MODELS IN GENETICS:
Insights into Students' Conceptions of Scientific Models

Abstract

Past research in a high school genetics classroom has shown that groups of students solving computer-generated genetics problems differed in their application of assessment criteria to the tentative models they generated and that consistent application of certain model-assessment criteria correlated with success in proposing adequate explanatory models. In this study in one high school class of 19 students, we focused classroom instruction on four specific model-assessment criteria: two involving fit between data and model, and two involving conceptual consistency (fit within the model itself or between it and other models or scientific knowledge). We monitored individual students' use of those assessment criteria during two rounds of model-revising problem solving. All students were able to assess models based on empirical fit, but, despite explicit instruction on their use, students did not consistently apply the conceptual assessment criteria to their models, perhaps in part because they lacked an understanding of models as conceptual rather than physical entities. In those instances when conceptual consistency criteria were applied to assess proposed models, students were able to identify constraints that directed their model-revising in fruitful ways. In this paper, we discuss the implications of these results on the use of modeling as an approach to teach students about the structure of scientific knowledge and the nature of science as a modeling activity.

Models are becoming an important topic of discussion in contemporary education. Researchers are weighing in on the pedagogical pros and cons of including models in science and math curricula, the psychological learning theories central to such model-based curricula, and pragmatic strategies for designing classrooms that enable students to learn in these often non-traditional environments. Although we find all this attention to scientific models exciting, it also generates some confusion. At the core of the confusion lie the many (often mutually exclusive) ways in which the term "model" is used. The particular view of "scientific" models that has shaped our work in curriculum design and research on student understanding in biology is one that takes into account what scientific practice actually entails—namely the construction and assessment of *explanatory* models.

The nature of scientific models. Perhaps the simplest way to begin defining what a scientific model is would be to point out what it is *not*. Educators as well as lay people often use the term "model" to describe (among other things) physical replicas of objects or systems. This use would include, for example, a space-filling molecular model made of plastic as well as the

material globes and light bulb that make up a "model" of the solar system. The term model is also used to refer to representational systems (e.g., maps, or diagrams) and mathematical algorithms or formulae (Lehrer & Schauble, in press). Not surprisingly, researchers characterizing middle and high school students' views on models have found that many students also cite examples of "models" that are physical replicas, verbal or visual entities, and mathematical formulae (Grosslight, Unger, Jay & Smith, 1991).

We recognize that these types of entities, namely representations, formulae, and physical replicas, play important roles in the science and mathematics curriculum and are sometimes prerequisites to the formation of scientific models. However, we take the position that they are not models in themselves. In our view, one that has been informed by the science studies community (see Giere, 1988; Kitcher, 1984;), a scientific model is a set of *ideas* that describes a natural *process*. A "scientific" model so conceived can be mentally "run," given certain constraints, to explain or predict natural phenomena. Once constructed, models influence and constrain the kinds of questions scientists ask about the natural world and the types of evidence they seek in support of particular arguments.

In the discipline of genetics, practitioners construct explanatory models to account for a variety of inheritance phenomena. The philosopher Philip Kitcher (1984) described explanatory "schema" in genetics—analogous to what we are calling genetic models—as providing the following information:

> (a) Specification of the number of relevant loci and the number of alleles at each locus; (b) Specification of the relationships between genotypes and phenotypes; (c) Specification of the relations between genes and chromosomes, of facts about the transmission of chromosomes to gametes (for example, resolution of the question whether there is disruption of normal segregation) and about the details of zygote formation; (d) Assignment of genotypes to individuals in the pedigree. (p. 356)

In keeping with this view of genetic practice as a modeling activity, we developed a 9-week high school biology module in which students used a Mendelian model of simple dominance to make sense of inheritance patterns in fruit flies and then revised that model in response to anomalous data (Cartier & Stewart, in press; Johnson & Stewart, 1990). Groups of students used the Genetics Construction Kit (GCK) software (Calley & Jungck, 1997) to generate populations of hypothetical flies and attempted to account for different inheritance patterns such as codominance and linkage. Students' explanatory models specified the number of alleles at given loci and the genotype-to-phenotype mappings for populations of organisms (Hafner & Stewart, 1995).

Assessment of explanatory models. In practice, scientific models are continuously assessed on the basis of empirical and conceptual criteria (Darden, 1991; Laudan, 1977; Stewart and Rudolph, in press). Specifically, scientists assess whether a particular model can explain all of the data at hand and predict the results of future experiments (empirical assessment). More-

over, because scientific theories are constructed of *families* of models and their relations to real-world phenomena (Giere, 1988), scientists must also evaluate how well a given model fits with other accepted models and outside knowledge (conceptual asssessment). Models that fail to satisfy some or all of the assessment criteria are discarded or (more commonly) revised until they are deemed acceptable. Thus, models are continuously revised as they are used to probe new phenomena and collect additional data.

Kitcher (1984) described how models are used and assessed during inquiry in genetics:

> After showing that the genetic hypothesis is consistent with the data and constraints of the problem, the principles of cytology and the laws of probability are used to compute expected distributions of phenotypes from crosses. The expected distributions are then compared with those assigned in part (d) of the genetic hypothesis. (p. 356)

This is somewhat analogous to what our genetics students did when solving model-revising problems using GCK. Students regularly tested proposed models for fit with existing data, making sure that they could account for the phenotypes of offspring that resulted from particular crosses (matings of flies). Sometimes students also checked their models for predictive power and consistency with other models such as meiosis, although this check for conceptual consistency was a strategy that students used sporadically (see below).

Past research in the genetics classroom. Much of our research in this genetics classroom to date has focused on analyses of students' problem solving as they proposed and revised models to account for data (Finkel & Stewart, 1994; Hafner & Stewart, 1995; Stewart & Hafner, 1991; Stewart, Hafner, Johnson, & Finkel, 1992; Wynne, Stewart & Passmore, 1998). Recent studies have paid particular attention to the strategies students used to assess their proposed models (Johnson & Stewart, 1998; Wynne et al. 1998). Johnson and Stewart (1998) found that one important distinction between successful and unsuccessful model revisers centered around their application of model-assessment strategies to tentative models. In general, all student groups were able to propose tentative models during each model-revising exercise; however, the groups differed in their ability to tinker with those models and produce final models with which they were satisfied. The students' abilities to productively revise proposed models were closely tied to their model assessment skills. In particular, successful groups were more likely than unsuccessful ones to assess their models in relation to other models, such as meiosis.

In the study described in this paper, we attempted to help students learn to assess models more effectively by providing them with explicit instruction on assessment criteria (both empirical and conceptual). We chose to focus on a subset of assessment strategies that seemed applicable to scientific models in general:

> 1. Assessment for explanatory adequacy: Does the model account for all the data available?
>
> 2. Assessment for predictive power: Does the model successfully predict new data?

3. Assessment for fit with prior knowledge or ideas (external consistency): Is the model consistent with what the student already knows about biology or science in general? Is the model consistent with what the student believes about the way the world works?

4. Assessment for fit between ideas within the model (internal consistency): Do all the elements or assumptions of the model fit with one another without contradiction?

Our push to help students appreciate the conceptual issues surrounding model assessment in addition to the empirical ones stemmed from a belief that understanding in science encompasses both a facility with subject-matter knowledge and familiarity with ways that knowledge is generated and justified in scientific practice. In other words, to really know science, a student must develop an understanding for epistemological as well as content-specific aspects of the discipline. Thus, we felt our students' understanding of genetics would be incomplete without some sense of the ways explanatory models for inheritance patterns are evaluated and made to fit within a larger context of conceptual knowledge in genetics.

Research Design

PARTICIPANTS

The students who participated in this study were high school juniors and seniors enrolled in an upper-level elective science course. The high school enrolled approximately 500–600 students and served both suburban and rural communities surrounding a midsize midwestern city. The 19 students in the genetics class had a variety of career objectives, ranging from attendance at a 4-year college to immediate employment following high school graduation.

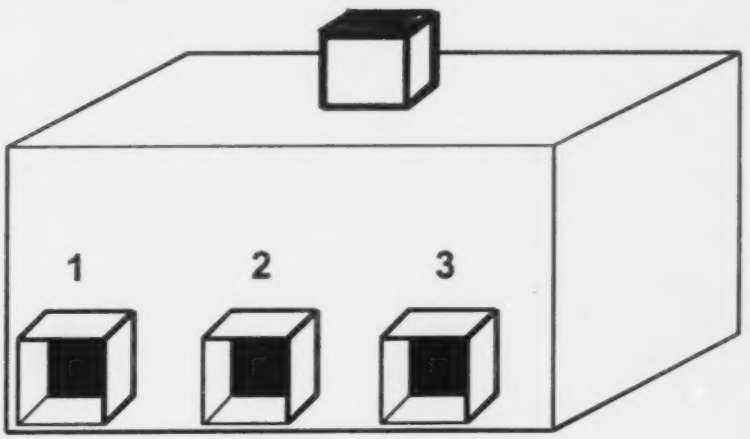
The teacher, a seasoned educational researcher herself, had 25 years of teaching experience and had been teaching the genetics course for nearly a decade. Development of the original course was part of her master's degree project, and subsequent research on student problem solving in the course became the focus of her doctoral dissertation.

ACTIVITIES

Black Box. During the first three days of the class, students were given a description of a box that had a single opening on top and three openings in the side and were asked to propose a model for phenomena associated with the box (see Figure 1 for student handout). Specifically, the students needed to explain why a white marble emerged from the box when a black one was put in (and vice versa) and why the marbles always emerged in succession from the three side openings. For the first day and a half, they worked in groups of two to four students, brainstorming ideas and drawing representations of their black box models on posterboard. Once

each group had finished its drawing, the groups presented their models to their classmates and responded to questions and criticism offered by their peers. The next day was devoted to revision of the models, and presentations of the final models followed on the third day.

Several students perceived the goal of the black box activity to be to correctly guess what was actually in the box, rather than to propose a plausible model with explanatory power. In



The following is a description of a "black box" that I saw at a conference. I wasn't able to get it, or build a copy. So, for the time being, this description will have to suffice.

Outside appearance:

The box happened to be an off-white, sort of beige color, as opposed to "black." It stood about 30 cm high, was approximately 50 cm wide and 30 cm deep. Located in the center on the top was a small, chimney-like opening, about 2 cm x 2 cm. On the front at the bottom were three openings similar to the opening on the top. These were numbered 1, 2, 3, from left to right. Nothing could really be seen inside the openings–it was mostly just dark. And there were no other openings or appendages to the box of any kind–no levers or switches on any side, no electrical cords, nothing. (See picture.)

Operation:

The person demonstrating the operation of the box had two containers of balls (about the size of large gumballs), one filled with black ones, the other with white. First a black ball was dropped into the opening on top of the box. After about 3–5 seconds, a white ball came out of the #1 opening on the front. Next, a white ball was dropped in the top. A black ball came out of opening #2. Throughout the time of the demonstration, whenever a white ball was put in, a black one would come out. And whenever a black ball was put in, a white one would come out. All the balls, regardless of color, always came out first chute 1, then chute 2, then chute 3, and then back to chute 1, and so forth. The order was always 1, 2, 3, 1, 2, 3 and always the emerging ball was the opposite color of the one most recently dropped into the top of the box.

FIGURE 1: STUDENT HANDOUT DESCRIBING THE APPEARANCE OF "BLACK BOX" AND SPECIFIC PHENOMENA ASSOCIATED WITH ITS OPERATION

BEST COPY AVAILABLE

fact, there was no physical box corresponding to the black box description they had received and so no such comparison could be made. Once the black box activity was completed, students were asked to write on the following questions in their journals: (a) What did the black box activity teach us about science? (b) How did I come up with my ideas [for building the black box model]? The teacher then spent one class period discussing with students their answers to these questions and creating a class list of modeling strategies that included—

Collect ideas from group members.

Brainstorm.

Rely on previous knowledge about black boxes and everyday things.

Create a representation.

Use an analogy.

Test for explanatory adequacy.

Test for predictive adequacy.

Students were told that they would revisit this modeling strategy list and add to it as the class proceeded. The discussion of models was then put aside, and two weeks of instruction on Mendel's model of simple dominance, meiosis, and use of GCK software ensued. Strategies for production and evaluation of models were discussed again only after the students had acquired facility with applying the simple dominance model to GCK problems.

Simple Dominance Model. Following the black box activity, the students read an edited version of Mendel's (1865/1959) paper *Experiments on Plant Hybridization* and were visited by a university professor playing the role of Mendel. Together with the students, "Mendel" examined sets of peas, noted the frequency of particular variations in each generation, and listed the results on the chalkboard. Then the students worked with the teacher to develop a model of simple dominance that they subsequently called the Mendel model (see Table 1). Next, working in their small groups, students spent several days in the computer lab using the Mendel model to explain inheritance patterns in computer-generated populations of hypothetical fruit flies. The objective was for the students to use Mendel's model of simple dominance to assign genotypes to each of the phenotypic variants in their populations. Once the students had acquired facility with one-trait problems, they were given problems in which they needed to explain the inheritance patterns in two traits—in other words, they needed to consider the underlying process of independent assortment when making assignments of genotypes to phenotypes in their populations.

Following work with the Mendel model, students received instruction on meiosis. They then returned to GCK problems and were asked to explain their data using a meiotic model. In order to do this, the students had to use Punnett squares to describe the possible gametes and fertilized eggs that would result from each cross they performed.

Model Revising. At this point, focus turned again to model-assessment criteria. Students were given a handout that described the four principle assessment strategies (explanatory power,

TABLE 1: MODELS OF INHERITANCE FOR A SINGLE TRAIT

Model	Genes per Trait	Alleles per Gene in the Population	Alleles per Gene in Individual	No. of Genotypes Possible	Genotypes			Dominance Relationships	Genotype / Phenotype Mapping				No. of Variations (phenotypes) possible
					(homozygous)	(heterozygous)	(single allele)		Variation A	Variation B	Variation C	Variation D	
Simple Dominance ("Mendel Model")	1	2	2	3	(1,1), (2,2)	(1,2)	--	1 is dominant to 2.	(1,1), (1,2)	(2,2)	--	--	2
Codominance	1	2	2	3	(1,1), (2,2)	(1,2)	--	1 and 2 are codominant	(1,1)	(2,2)	(1,2)	--	3
Multiple Alleles	1	3	2	6	(1,1), (2,2), (3,3)	(1,2), (1,3), (2,3)	--	Each pair examined: (1,3), 1 is dominant to 3; (2,3), 2 is dominant to 3; (1,2), 1 and 2 are codominant.	(1,1), (1,3)	(2,2), (2,3)	(1,2)	(3,3)	4*
X-Linkage	1	2	2 (F), 1(M)	5	F: (1,1), (2,2)	F: (1,2)	M: (1,-), (2,-)	1 is dominant to 2.	F: (1,1), (1,2) M: (1,-)	F: (2,2) M: (2,-)	--	--	2

* Our example results in four variations, but as many as six variations are possible. If no allele is dominant to any other, the three homozygous combinations — (1,2), (2,2), (3,3) — result in three unique variations; the three heterozygous combinations — (1,2), (1,3), (2,3) — result in three different variations.

BEST COPY AVAILABLE

predictive adequacy, external and internal consistency) and a hypothetical GCK scenario in which a student proposed a model inconsistent with meiosis. (Specifically, the model was inconsistent with the processes of both segregation and independent assortment.) The students were asked to evaluate the model in terms of the assessment criteria listed on the handout and discuss their ideas with the class.

Finally, the students were set to work on GCK problems in which traits were inherited in a non-Mendelian (i.e., a non-simple dominant) fashion. Consequently, students were not able to use the Mendel model to explain the inheritance patterns they saw. They needed to alter various components of the model in an attempt to produce a new model with which they were satisfied. During the first round of model revising, each group of students worked to explain the pattern in a trait inherited in a codominant fashion. During the second round of model revising, some of the groups worked on a multiple-alleles inheritance pattern while other groups worked on X-linkage inheritance (see Table 1 for summaries of these models). Each round of model revision took three to four days and culminated in a class conference during which students presented their models and answered questions posed by their peers. The teacher requested that students present even those models that were still in the formative stage and asked that all students discuss some of the models they considered but subsequently discarded. She encouraged students to talk about why they liked and disliked various models and to use the vocabulary of model assessment (i.e., "explain," "predict," "consistent") during these discussions.

Throughout the class, students recorded their work in research notebooks and wrote journal entries in response to homework questions. Two exams were given, one at midterm and one at the end of the quarter.

INTERVIEWS

Eight students (3 boys, 5 girls) were chosen to participate in a series of open-ended interviews designed to elicit students' strategy use as they revised their inheritance models. Choice of participants was based solely on the students' schedules and availability: Because we wanted to conduct the interviews outside of class time, but during the school day, only students with at least one study period were chosen. The interviews, lasting 5–20 minutes each, were conducted after the introductory black box activity, during each day of model revising (rounds one and two), and following the final class discussion of models for multiple alleles and X-linkage. Each student was interviewed 6–8 times. Interviews were audiotaped and transcribed. Students were asked initial questions about the status of their work with GCK, such as "Are you working with a model right now to explain this new inheritance pattern? Can you tell me about it?" Follow-up questions were aimed at elucidating both students' current models and previous models that they had considered and rejected. Attention was paid to the reasons students gave for rejecting models or for being satisfied with final models. The researcher studied records (both computer-generated records of GCK data and student-generated logs of their work) each evening prior to the interviews and based individual questions on each student's own work. For instance, a student might be asked, "Yesterday we talked about a three-allele model, but today I noticed that

you seem to have stopped using that model. Can you tell me about that choice?" In addition, the initial and second interviews included one question about the nature of models in general: Students were asked, "Can you tell me some of the characteristics of a good model? Conversely, what are some of the characteristics of a model you would consider to be bad or not as good?" In the final interview, students who had proposed a second-round model with which they were satisfied were shown representations of two competing models for the same inheritance phenomenon and asked to identify which model they liked best. They were then asked to explain the reasons for their choices.

CODING

Transcripts of interviews were coded for use of empirical and conceptual model assessment strategies. In particular, aspects of Darden's (1991) and Laudan's (1977) frameworks for thinking about problems and problem solving in genetics provided the following initial codes:

1. assessment for explanatory adequacy.
2. assessment for predictive adequacy.
3. assessment for internal consistency.
4. assessment for external consistency.

Additional codes emerged from the data as students' assessment criteria were studied and discussed. These more pragmatic strategies included (a) assessment for simplicity; (b) assessment for ease with which the model could be used; (c) assessment for a match with underlying reality; and (d) assessment of the quality of the representation of the model. Once instances of students' use of assessment strategies had been identified, it was noted that students made use of conceptual criteria for assessing models less frequently than they did of empirical criteria. Consequently, the transcripts were coded again to identify instances where students failed to use conceptual criteria to judge models—instances where conceptual inconsistencies were present but overlooked or deliberately ignored. Finally, we used the classificatory framework of Grosslight et al. (1991) to code the interviews for students' ideas about models. In particular, we identified instances where students referred or alluded to models as being (a) objects, (b) visual entities, (c) verbal entities or explanations, or (d) abstract ideas or concepts.

Results and Discussion

In this paper, we argue that students assessed explanatory models in genetics mainly based on empirical criteria. Occasionally, they also judged the conceptual adequacy of models based on how well those models fit with what the students already knew about science (external consistency) or on the degree to which the elements of a single model fit with one another (internal consistency). We illustrate our claims with examples of students' problem solving and transcripts of interviews in which students discussed their tentative explanations as well as their reasons for discarding or retaining particular models. We also examine interview data for

evidence of students' understanding of the nature of models in genetics. Finally, we discuss the implications of our results on the use of modeling as a mechanism to help students understand both conceptual and epistemological aspects of genetic practice.

HOW STUDENTS ASSESS EXPLANATORY MODELS FOR INHERITANCE PHENOMENA

During both rounds of model revising, each of the eight students interviewed assessed their tentative and final models based on the degree to which those models could explain their data. Lack of explanatory power was the most common reason that students discarded a tentative model. For example, one student was working on a model for codominance, assigning genotypes to the individuals in each of the crosses he had made. In his original model, he suggested that two of the variations were recessive phenotypes (bithorax and tetraltera) and the other was the dominant phenotype (grooveless):[1]

Interviewer: OK, so you said that you started with a model that was more like you had two recessives and one dominant? What did that model look like? What were the recessives?

Connor: The recessives was bithorax and I think we thought tetraltera was the other one. But we got rid of that idea pretty quickly.

Interviewer: Why did you get rid of that idea?

Connor: Because after we started making more crosses, it just didn't mix. I guess you might say.

Interviewer: It didn't mix?

Connor: It didn't work.

Interviewer: What do you mean it didn't work?

Connor: For tetraltera to be recessive. I don't know how to say that.

Interviewer: OK, so you made some crosses, and you decided that your model didn't work. What was it that you were seeing that made you decide that your model didn't work?

Connor: I thought it was one of the vials [computer-generated vials of flies—the result of a cross] here, but I'm not seeing it. Um [pause] she [another member of the research group] really doesn't have it written out. I thought we had it on a vial, but maybe not. Well, wait, what was your question again?

Interviewer: I'm just wondering why you decided to get rid of your first model where you thought that both bithorax and tetraltera were recessive. And grooveless then, I guess, would have been dominant?

[1] Note that this student is using the terms "recessive" and "dominant" to describe phenotypes rather than to describe alleles, as in Mendel's simple dominance model. This inconsistent use of vocabulary was quite common once students began working with inheritance patterns that did not fit their original notion of simple dominance. Most commonly, students identified a phenotype or variant as being dominant if it resulted from the underlying genotype (1,1). Similarly, a variant was labeled recessive if it resulted from a genotype with at least one '2' allele in it (i.e. 1,2 or 2,2). Consequently, students frequently labeled more than one variant in a population as recessive.

BEST COPY AVAILABLE

Connor: OK, well, this kind of helps it. Vial 9. We had a bithorax and a tetraltera and we got all grooveless. That means that we're taking all 1,1 [a genotype he is assigning to the dominant variation, grooveless] from a 2,2 [a genotype he is assigning to both recessive variations]. We couldn't have a 2,2 and a 2,2. If that makes any sense.

Interviewer: Right, so I guess what you're telling me is that when you thought both bithorax and tetraltera were recessive, you thought they were both 2,2's?

Connor: Yeah.

Interviewer: OK, and then when you do a 2,2 x 2,2 cross and you get all grooveless—

Connor: That kind of proves you wrong. It shows you that one of them [one of the parents–either the tetraltera or the bithorax] had to be a 1,1.

Here Connor rejected his tentative model based on its inability to account for the data produced in his ninth cross. He later revised his model and based his satisfaction with the new model on how well the data ("the vials") supported it:

Interviewer: OK, so tell me more about your model because we sort of got off the track here. So you told me that your new model now is that you have two dominants, tetraltera [(1,1)] and grooveless [(1,2)], and one recessive, bithorax [(2,2)].

Connor: What do you want me to talk about with them? I mean, you already know what they are.

Interviewer: Um [pause] do you think this is a good model?

Connor: Yes. I think it is. I think it backs itself up in the notes, too.

Interviewer: What do you mean by that?

Connor: If you use a little bit of the old model, the Mendel model, and you look at ours [pause] you look at the vials. You look at the crosses, this kind of proves itself, I guess. I guess it just shows that it works. I'm not sure how else to explain it.

The ability to explain cross results was the most common reason students gave for being satisfied with a model. Moreover, when asked about the advice they would give to other students trying to decide whether or not to keep or discard a tentative model, most students suggested that any model that could explain the data at hand should be kept:

Interviewer: What if this person already had a model and she wanted to know whether she should keep revising it or whether she should stick with her model? What would you tell her?

Mac: Try it out. I mean, if it works for all the vials, then it's a good model. As far as we have seen. I mean, all the models that explain, like, every cross are good, seem to be good models so far. Some are more complicated than others, but they seem to make sense.

[...]

Griffin: If it works for every cross, keep it. If it only works for [certain] things like ours does, try to get a better one.

[...]

Jenner: Actually, just try to do crosses and if they keep working, the chances of your model being right are greater. If you can keep explaining what you get.

Other students extended their empirical criteria beyond the ability to explain extant data and found satisfaction with a model when it could be used to correctly predict additional data. Examples of this are found in two other students' responses to the question above:

Cassidy: It's always good to have other people's input. Like when you guys handed out Amy and Connor's, like even Elizabeth's [representations of their models]. You know. See if it works for other people, too.

Interviewer: So test your model on other people's data you mean?

Cassidy: Uh huh.

Interviewer: How would you use your model to test other people's data?

Cassidy: You mean, like, if they give us a problem like this? Just start crossing them, and see if we come up with the right, or the probable, results.

Interviewer: So you would do a cross, and see if your model could explain that cross, or you would predict what you would get first and then—

Cassidy: We did a lot of that this time. Predict. Or Ms. Lambert, when she was with us, she would ask first before we crossed it. I guess, what if you already have a model done, then you would predict. Because you'd cross expecting a certain thing and then if you didn't get it, you would know that it didn't, wouldn't work.

[...]

Hannah: Just do the Punnett squares with it [construct a Punnett square diagram with possible gametes and fertilized eggs, given the parental genotypes]. If she's got at least that far.

[Discussion continues]

Hannah: Fill out the Punnett squares like what we thought the next cross would be and see if it came out, and then if it did then we knew that our model was working.

Although students were usually quite systematic about examining proposed models for their ability to correctly explain and predict data, they were less so about examining models for conceptual consistency. However, some students did reject models based upon *internal* consistency issues (where elements of a model were inconsistent with one another or a model was inconsistently defined across a single set of data), such as in the following example:

Interviewer: OK. Are you guys working with a model now [to explain the codominance inheritance pattern]?

Griffin: Um [pause] we don't get it.

Interviewer: Ha Ha! That's not what I asked you! Do you have any ideas at all about how you're going to explain this data?

Griffin: We don't [pause] I don't know. We can't figure anything out. 'Cause we say we took a 1,2,3 and a 1,2,3 [crossed two individuals with genotypes (1,2,3) and (1,2,3)], and the kids still end up being two things.

Interviewer: Still end up being two things?

Griffin: Two like genotypes [two alleles per genotype]. Or like a 1,1 and those other ones are 1,2,3 and the kids are 1,1. She said that doesn't work.

Interviewer: Ms. Lambert said that doesn't work?

Griffin: I can't remember. It was one of you guys.

Interviewer: Do you see that there might be a problem with that model?

Griffin: Yeah the parents [pause] the kids have to be the same genotype. Well, not the same genotype, but, like, the same amount. Same number of genotypes [same number of alleles per genotype].

Interviewer: Same number of alleles you mean?

Griffin: Yeah.

In this example, Griffin was dissatisfied with his proposed model because the number of alleles per individual (or genotype) fluctuated, or was inconsistent, from generation to generation. Perhaps more interesting, however, is the fact that Griffin failed to acknowledge the external inconsistencies associated with the model, that is, his model was inconsistent with what the students had been taught about the meiotic model--the model that defines the underlying processes of segregation and assortment. According to the meiotic model, it would be impossible for three alleles for a given gene (the number of alleles the parents had in Griffin's model) to segregate evenly into gametes. Moreover, the meiotic model specifies that each gamete has exactly half the number of genes as the parent had. In fertilization, two gametes combine to restore that normal complement of genes.

The significance of the above example lies in its prevalence throughout students' problem solving--a number of students proposed models in which each individual had an odd number of alleles for a given trait. Most students discarded these models based on their lack of internal consistency (as Griffin did) or their lack of explanatory power. For example, Libby's group initially proposed a single-allele-per-individual model to explain codominance but rejected the model based on their inability to explain their cross results. Next, her group suggested a three-allele-per-individual model identical to Griffin's and again rejected the model based on its lack of explanatory power. When asked why they discarded these models, neither Libby nor her groupmate Jenner mentioned the inconsistencies these tentative models had with meiosis. Instead, they talked only about their problems making the model work with their cross data.

In this first round of model revising, neither Griffin's nor Libby and Jenner's group was able to propose a codominance-inheritance model with which they were satisfied. We suggest that attention to the nature of the meiotic inconsistencies within their models might have pointed

out viable alternatives to these students. We base this suggestion in part on data from a later class when students working in larger research teams also pointed out the problem of internal inconsistency within a model that postulated three alleles in the parental generation and only two in the offspring. However, this group of students probed further and hypothesized about scenarios that would have to occur during gamete formation in order for their model to be possible. As this discussion was occurring, a number of students were grouped around the chalkboard. Some were writing cross results or making genotype assignments for particular crosses. Students not at the board were involved in the discussion as well, calling out genotype assignments, summarizing the models that were being tested, and taking notes. The students were testing models to explain the inheritance of a trait with five variations (a multiple-alleles pattern). After about thirty minutes during which the students proposed and tested two models, the teacher entered the room and noticed two sets of genotype assignments on the board. One set was from a three-allele-per-individual model, and the other set from their final three-alleles-per-population (but two-per-individual) model. She prompted the students to explain why they had discarded the former model:

Ms. Lambert: I'm confused. I'm just curious. I'm a newcomer to this research lab. I see you using two numbers in some areas and three numbers in other areas. [For example, (1,2,2) was a genotype in their three-allele-per-individual model and (1,2) was a genotype in their three-allele-per-population model.]

Kelly: We forgot about that three-number trait.

Sarah: Cross that out. It didn't work.

Kelly: We didn't know how we would cross something that had three alleles and another thing that had three alleles to get a kid that had three alleles.

Anne: The Punnett squares we did just didn't want to do that.

Sarah: We had to stick with only two alleles [per individual], so then we just made it three different kinds of alleles [in the population].

Jill: Or else you just get, like, two from one parent and then one from another. Or else three from another.

Earlier in their discussion, the students had recognized that the problem with segregation could be solved only by postulating models with an even number of alleles per individual. However, they knew they needed more than two different alleles to generate enough variety in genotypes to account for the five different phenotypes they had observed. These constraints led them to suggest three different alleles in the population, but only two alleles per individual, a model that eventually worked well to explain their data and with which they were ultimately satisfied.

Inattention to consistency with meiosis occurred in several modeling contexts and did not always (as it did for Connor and Libby) result in models that students found dissatisfying. In fact, students sometimes overlooked or deliberately ignored such inconsistencies within models they judged to have adequate explanatory or predictive power. For example, Cassidy and her group

worked on a model to explain a multiple-alleles inheritance phenomenon (one in which there were three alleles and four variations). Their initial group model described the following phenotype to genotype matches:

variation A (abero) (1,1)

variation B (cut) (1,2)

variation C (lyra) (2,2)

variation D (bobbed) (3,3)

When asked about her satisfaction with this model, Cassidy said it was a good model because "everything worked about it" and that "just [by] making crosses and seeing if the right variation—the probable results—showed up" she and her group members were able to determine that it was an adequate model. Thus, she judged the model based upon empirical criteria—how well it matched her predictions of future cross data. Her assessment of the model did not include a check for consistency with meiosis. Instead, she overlooked the fact that the model did not postulate the existence of individuals with genotypes (1,3) or (2,3), and that these individuals had to exist (or at least be possible) if the model was consistent with the meiotic process of segregation and subsequent fertilization. It was only after her group performed a cross that they were unable to use their model to explain that they acknowledged the need for (1,3) and (2,3) individuals and revised their model accordingly. It is noteworthy that Cassidy never acknowledged any conceptual implications that led up to or arose from the inclusion of (1,3) and (2,3) individuals in her model. Rather, these genotypes were matched to phenotypes within the model solely because of their empirical utility.

A different group of students, working on a model to explain X-linkage, did make use of the meiotic model and their knowledge about sex determination to assess their initial model and to constrain their final one. Specifically, their model postulated a third allele (3) that was always associated with maleness: Males always had one 3-allele and females never did. Hannah described how their group decided that only males could have a single 3-allele:

Hannah: Yeah. So that's how we figured out that it had to be something male and female. And then we decided to keep the same Mendel genotypes [preserve the simple dominance relationship between alleles 1 and 2]. And, but we'd have to come up with different ones [pause] we decided to keep the same ones for the female [(1,1); (1,2); (2,2)] but come up with different ones for the male. And we brought in a third allele which is the 3. And then we did have, you can't see it on here, but we did have, like, a 3,3. But we figured that [pause] we decided to cross it out because there's no way that you could ever get a 3,3. So we just have two [genotypes–(1,3); (2,3)] for the male now and three for the female.

Interviewer: Why couldn't you ever get a 3,3?

Hannah: Because you'd need a 3 from the female too. You'd have to start out [pause] cause that's how we can identify if it's a male. If it has a 3 in it. And if it doesn't have a 3 in it, then it's a female.

Interviewer: OK. So every male has to have a 3.

Hannah: Yeah.

Hannah was then pressed further to be explicit about how well her model fit with her understanding of meiosis and sex determination. Specifically, the researcher wanted to know whether Hannah could explain how the 3-allele always became associated with maleness:

Interviewer: I just wanted to know what you guys are thinking about in terms of what this third allele is. And you just told me that every male has to have a 3, but no females can have a 3. So now I want to know, with what you know about meiosis, how is that third allele working?

Hannah: Well isn't it kinda just like the Y? With the X and Y?

Interviewer: Can you say more about that?

Hannah: [inaudible] cause I don't really [pause] I don't really understand. We never discussed that as a group so I don't really know how it would work with—

Interviewer: So what do you know about what the Y chromosome does? You're saying that it's kinda like the Y chromosome?

Hannah: That's what the 3 is, yeah.

Interviewer: That's what the 3 is. OK

Hannah: So only the males carry it, and they pass it down. In the males.

Another group, working on the same problem, was unable to propose a model with which they were satisfied. Ultimately, their model used two alleles (1 and 2) and two subscripts (a and b) to describe the inheritance pattern for X-linkage. Their model was able to explain some, but not all, of their data and had essentially no predictive power. For these reasons, the group was dissatisfied with the model. Interestingly, their model postulated that (aa) individuals would be female, (bb) individuals male, and (ab) individuals could be either male or female. When asked how maleness or femaleness was determined in the (ab) individuals, Connor responded that it might be "due to chance variation" and admitted that he "really [couldn't] figure that out at all." His inability to explain this aspect of his model did not seem to affect his satisfaction with it. Rather, it was simply the lack of predictive power associated with the model that Connor identified as being problematic. This serves as another example of students assessing a model solely on empirical criteria and ignoring conceptual inconsistencies between the model and other biological knowledge (in this case, sex determination). Given the success of Hannah's group, it seems likely that paying more attention to the details of sex determination and focusing the problem on how that knowledge fit with their model might have constrained Connor's group's problem solving and pointed out new ideas for them to incorporate into later models.

Summary. Our results indicate that students assessed explanatory models for inheritance patterns mainly on empirical criteria. Even when conceptual inconsistencies occurred between their proposed models and other models or biological knowledge, their primary focus was usually on how well a given model could explain the data at hand. The principal exception was when internal inconsistencies occurred: These were usually recognized by students and led to

BEST COPY AVAILABLE

immediate dissatisfaction with the model(s) in question. In contrast, students frequently had difficulty recognizing specific inconsistencies between their models and meiotic concepts or other biological knowledge such as the method of sex determination in humans or fruit flies. In some instances, students recognized that their models were inconsistent with other knowledge but were willing to overlook such inconsistencies when their models were judged to have adequate explanatory power. Thus, students paid more attention to empirical than to conceptual issues and tended to value empirical power over conceptual consistency in models where both criteria were brought to bear.

Compared with the rest of the class, those students who paid particular attention to conceptual consistency (mostly with meiosis) were more likely to propose models with which they were satisfied and which were able to explain all of their data. Moreover, the identification of particular consistency issues led to the recognition of modeling constraints and directed model revision in fruitful ways. For instance, students realized that in order for models to be consistent with the processes of segregation and fertilization, an even number of alleles (or at least chromosomes) needed to be postulated for each parent; students also recognized that an allele that defined "maleness" (the 3-allele in the X-linkage model) could only be present in a single copy in male organisms.

IMPLICATIONS FOR STUDENTS' UNDERSTANDING OF GENETICS

Given these observations, it might be easy to suggest that attention to conceptual issues could aid students in proposing explanatory models in genetics and be satisfied with that suggestion. However, it is not our aim to help students propose successful models in genetics per se, but rather to help them build an understanding of genetics in general. Such an understanding includes facility with some conceptual knowledge of genetics as well as familiarity with epistemological aspects of knowledge construction and justification in the discipline. Our results suggest that using modeling as a framework for teaching genetics could provide rich opportunities to introduce students to the structure of genetic practice.

Borrowing from Giere's (1988) notion of the structure of scientific theories, we suggest that students should be encouraged to think of genetics as consisting of a "family" of interrelated models. In this particular class, students might be led to consider simple dominance, meiosis, multiple alleles, X-linkage, and codominance as a family of models that share certain assumptions and constraints and together provide explanations for a broad array of inheritance phenomena. It is not enough for students to consider each model individually–as geneticists must always judge the worth of their work in relation to other accepted conceptual knowledge and methodological norms within their discipline. For example, Kitcher (1984) noted that "constraints" on geneticists' explanations for hereditary phenomena include—

> general cytological information and descriptions of the chromosomal constitution of members of the species. The former will include the thesis that genes are (almost always) chromosomal segments and the principles that govern meiosis. (p. 355)

Students whose model-revising experiences are focused on recognizing and valuing such consistency between various knowledge and conceptual models about inheritance should develop a more complete view of genetic practice than those who take a less macroscopic view. In addition, encouraging students to focus on how well various aspects of their biological knowledge fit with one another pushes them to articulate their ideas in a variety of contexts. It seems that this can only help students to firm up their understanding of biological concepts.

STUDENTS' CONCEPTIONS OF MODELS

Others have advocated a similar pedagogical approach to the one we are describing here, namely helping students learn about scientific inquiry by providing them experiences constructing and revising scientific models (Grosslight et al. 1991; White & Frederiksen, 1998). In their study of students' and experts' understandings of models in science, Grosslight et al. (1991) suggested that in order to facilitate students' appreciation for models as conceptual tools that guide scientific inquiry, educators need to—

> provide students with experiences using models to solve intellectual problems. In this way students would have the opportunity to learn that a model can be used as a tool of inquiry and that it is not simply a package of facts about the world that needs to be memorized.
>
> Second, students need experiences with multiple models of the same phenomenon and with revising and/or modifying models as they encounter new phenomena . . . In this way students may come to realize what a conceptual vantage point is and how it can influence one's thinking.

Thus, it seems that a curriculum focused on modeling tasks and issues such as the one in our genetics class could afford students many opportunities to grapple with the conceptual nature of models in genetics. In practice, however, we have found that actually getting students to engage these ideas is not straightforward. Pushing students to articulate their reasons for model choice, for instance, revealed that many of them made such choices based on distinctly pragmatic criteria: They liked the way one model was represented more than another; one model might have seemed easier to use than another; one model had fewer components to memorize than another, and so on.

Using Grosslight's (1991) framework for classifying students' understanding of models, we coded portions of our interviews to get a general sense of the ways in which our students were thinking about models. Specifically, after they had spent several days proposing and critiquing models to explain some black box phenomena, we asked the students what they felt were the important characteristics of a good model. We posed the same question to them again after they had practiced using the Mendel model to explain one- and two-trait GCK problems and had begun to revise that model to explain codominance. These portions of the interview transcripts were examined for evidence of students' thinking about models based on the following categories: (a) physical objects, (b) visual entities (such as pictures or drawings), (c) verbal entities (such as explanations of drawings), or (d) abstract entities or ideas.

Generally, students judged models based on how well they were represented (visual) or how well those representations were explained (verbal). For instance, the following are students' descriptions of good models:

(visual)

> "Um, like a good drawing of it." Libby
>
> "I consider a model [to be] physical things. Like the pictures we drew [to represent, the inner workings of the black box]. And on the computer it's easy for me to see cause I can, like, see how I crossed them and which ones I crossed and it's clear to me." Cassidy
>
> "The finer points of a model. Like this [a diagram of the inner workings of the black box] just shows overall the general thing. I'd like to show more depth of one area. It's kind of like on a map where they'll show, like, a city in more depth, like, in another part." Connor

(verbal)

> "The person making the model should be able to explain the details well." Connor
>
> "I guess I mean, just going on whether it's good or bad, I guess how they back up their conclusions. Cause anything's possible. They just have to explain why they think that." Mac
>
> "Um, you can explain what you wrote down. They have a lot of detail." Delia
>
> "Just have good—know what you did. Have good knowledge about your model. Like, be able to answer the questions that people ask about it." Libby

Of the eight students interviewed, all of them made statements characterizing good models according to aspects of the visual or verbal properties of their representations. Only one student (Hannah) mentioned that a good model is one that can be "applied" to explain data, consistent with our notion of models as ideas or abstract entities.

Additionally, several students described good models as those that were "simple" or "reasonable." When pressed to say more, they revealed that the quality of simplicity was desirable because the black box was small, and only simple things would fit within it. In other words, these students believed that the model in question was an object: Either an actual physical replica of the inner mechanism of the black box or a blueprint that could be used to construct such a replica. Thus, a model judged to be "reasonable" necessarily was constrained by the belief that there was a single real entity within the box and/or by the box's physical dimensions. For example, Mac stated, "I guess it's hard to say what would be a poor model 'cause nobody really knows what's in the black box, so it's hard to say if it's poor." Two other students discussed

models in the following terms:

Connor: First, like, come up with ideas about how it works, I guess. Pretty much just brainstorm for a while, and then take the most possible one—the most possible answer I guess you might say.

Interviewer: How do you decide which one is the most possible?

Conner: The one that looks and sounds more realistic than the others. Sounds like it could really happen. Or be made.

[...]

Interviewer: My last question is, what kinds of characteristics make up a good model?

Anne: I guess reasonable answers. Something that everybody would think is possible. Maybe. Like, there's just a lot of things that you just look at and you say there's no way that could even be possible because of this or 'cause of that. I think that could—they're the most realistic.

Interviewer: What do you mean by this or that? How do you decide whether something's possible?

Anne: Like, that magnet stuff that—actually not the magnet stuff [a model in which students proposed that the black and white balls were separated based upon their different magnetic polarities]. Like one person did something about a—something about the balls falling on weights inside and that telling them if it goes one place or another. I think that seems a little unreasonable.

Interviewer: What makes it unreasonable?

Anne: The box isn't big enough. There's too many little details. The stuff would be too expensive. The stuff wouldn't be big enough.

Conclusions. The difficulty students had confronting conceptual issues while model revising in genetics seems less surprising when we also consider their tendency to think of models as physical instantiations of ideas rather than as ideas themselves. Considering these issues together, we suggest that a great deal more attention needs to be paid to students' understandings of the nature of models at the outset of a class—specifically, students should be encouraged to think about families of genetic models that share particular conceptual elements and are embedded within a particular epistemological context that governs their formation and use.

If we, or any other educators, are to achieve the important goal of moving away from thinking about models as "[packages] of facts about the world that need to be memorized" (Grosslight et al., 1991) toward a deeper understanding of the epistemological issues surrounding the creation and justification of models, we need to find ways for students to engage such issues in real classrooms. Toward this end, our curriculum design collaborative has modified the genetics curriculum described herein and undertaken research into students' understanding of the nature of models in genetics as well as evolutionary biology (see Cartier, Stewart, & Johnson, 1998 for a description of particular genetics curriculum changes). In addition, we have continued our characterization of the criteria by which students assess models in these biological disciplines. Reports of these projects are forthcoming.

References

Calley J., & Jungck, J.R. (1997). Genetics construction kit (The BioQUEST Library IV, version 1.1B3) [Computer software]. College Park, Maryland: The ePress Project.

Cartier, J. & Stewart, J. (in press). Design and implementation of an inquiry-focused genetics curriculum for high school biology students. *Science and Education.*

Darden, L. (1991). *Theory change in science: Strategies from Mendelian genetics.* New York: Oxford University Press.

Finkel, E. A., & Stewart, J. (1994). Strategies for model-revision in a high school genetics classroom. *Mind, Culture, and Activity: An International Journal, 1,* 168–195.

Giere, R. N. (1988). *Explaining science: A cognitive approach.* Chicago: The University of Chicago Press.

Grosslight, L., Unger, C., Jay, E., & Smith, C. L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching, 28,* 799–822.

Hafner, R., & Stewart, J. (1995). Revising explanatory models to accommodate anomalous genetic phenomena: Problem solving in the "context of discovery." *Science Education, 79,* 111–146.

Johnson, S. K., & Stewart, J. (1998, April). *Revising and assessing explanatory models in genetics: A comparison of successful and less successful performance by high school students.* Paper presented at the annual conference of the National Association for Research in Science Teaching, San Diego, CA.

Johnson, S. K., & Stewart, J. (1990). Using philosophy of science in curriculum development: An example from high school genetics. *International Journal of Science Education, 12,* 297–307.

Kitcher, P. (1984). 1953 and all that. A tale of two sciences. *The Philosophical Review, 93,* 335–373.

Laudan, L. (1977). *Progress and its problems: Towards a theory of scientific growth.* Berkeley, California: University of California Press.

Lehrer, R., & Schauble, L. (in press). Modeling in mathematics and science. In R. Glaser (Ed.), *Advances in instructional psychology.* Mahwah, NJ: Erlbaum.

Mendel, G. (1959). Experiments on plant hybridization. In J. Peters (Ed.), *Classic papers in genetics.* New Jersey: Prentice Hall. (Original work published 1865)

BEST COPY AVAILABLE

Rudolph, J. (1998, April). *Reconsidering the nature of science as a curriculum component.* Paper presented at the annual conference of the National Association for Research in Science Teaching, San Diego, CA.

Stewart, J., & Hafner, R. (1991). Extending the concept of "problem" in problem-solving research. *Science Education, 75,* 105–120.

Stewart, J., Hafner, R., Johnson, S., & Finkel, E. (1992). Science as model building: Computers and high school genetics. *Educational Psychologist, 27,* 317–336.

Stewart, J. & Rudolph, J. (in press). Considering the nature of scientific problems when designing science curricula. *Science Education.*

White, B.Y., & Frederiksen, J.R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16.* 3–118.

Wynne, C., Stewart, J., & Passmore, C. (1998, April). *High school students' revising of meiotic models.* Paper presented at the annual conference of the National Association for Research in Science Teaching, San Diego, CA.